JOURNAL OF CLINICAL MICROBIOLOGY, Nov. 2002, p. 4021–4029 Vol. 40, No. 11
0095-1137/02/$04.00+0 DOI: 10.1128/JCM.40.11.4021–4029.2002

Genetic Characterization and Sequence Heterogeneity of a Canadian Isolate of Swine Hepatitis E Virus

Yanlong Pei and Dongwan Yoo*

Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received 6 June 2002/Returned for modification 3 July 2002/Accepted 14 August 2002

Swine hepatitis E virus (HEV) is a newly identified potentially zoonotic agent that is possibly transmitted to humans from pigs. Swine HEV is prevalent in pig populations and does not cause abnormal clinical symptoms in infected pigs, further implicating a likelihood of a risk of transmission to humans by normal contact. To date in North America, only one strain of swine HEV (strain US swine) has been fully sequenced. In the present study, we identified a swine HEV isolate from pigs in Canada, designated the Arkell strain, and determined the full length of the genomic sequence. The genome of Canadian strain Arkell consisted of 7,242 nucleotides, excluding the poly(A) tail of at least 15 A residues. The genome contained three open reading frames (ORFs), ORF1, ORF2, and ORF3, which had coding capacities for proteins of 1,708, 660, and 122 amino acids, respectively. Comparative analysis of the full-length genomic sequence indicated that the sequence of strain Arkell was distinct from those of all other known HEV isolates by 13 to 27% and shared the highest degrees of identity with human HEV isolates US-1 and US-2, HEV isolate US swine, and the human and swine HEV isolates recently isolated in Japan. On the basis of sequence similarities and phylogenetic analyses, HEV strain Arkell was grouped into genotype 3. The sequence of the Arkell swine HEV isolate differed from those of HEV isolate US swine and HEV isolate Japan swine by 13 and 14%, respectively. To date, two isolates of swine HEV (isolates Arkell and SK3 [D. Yoo et al., Clin. Diagn. Lab. Immunol. 8:1213-1219, 2001]) have been identified in Canadian pigs, and their sequences also differ from each other by 11.8%. Our studies indicate that, as with human HEV strains, swine HEV isolates exhibit extensive genetic heterogeneity.

Hepatitis E is one of the six recognized types (types A, B, C, D, E, and G) of viral hepatitis in humans. Among the six types, hepatitis E is known as food-borne, waterborne hepatitis, as it is believed to be transmitted primarily by the fecal-oral route via contaminated food or water. The clinical features of hepatitis E include jaundice, anorexia, nausea, and hepatomegaly. The mortality rate among those with hepatitis E is reported to be 1 to 3%, whereas it is 0.3% among those with hepatitis A; but unlike hepatitis B and C, recovery is almost always complete, without progression to a chronic state. Hepatitis E has been reported to be severe in pregnant women, with high rates of fulminating hepatitis and a case fatality of up to 20% (for reviews, see references 20 and 27). A recent study indicates that the mother-to-infant transmission rate is nearly 100% (12), suggesting congenital infection with hepatitis E virus (HEV).

HEV, the etiological agent of hepatitis E, is a small nonenveloped virus with a single-stranded positive-sense RNA genome of approximately 7.2 kb (9, 26, 29). The virus was initially considered a member of the family *Caliciviridae*, but on the basis of phylogenetic analyses of the genomic sequences, it was recently removed from that family and was reassigned to an unclassified genus termed "hepatitis E-like virus" (2). HEV is endemic in areas of Asia, South America, the Middle East, and Northern and Western Africa. Genomic similarities and phylogenetic studies of viruses isolated from these countries indicate that HEV strains can be grouped into two distinct genotypes: the Asian genotype and the Mexican genotype. In contrast, HEV has been considered absent from countries in Western Europe, Oceania, and North America (the United States and Canada), and the clinical cases of hepatitis E reported in these countries have primarily been known to be associated with travel to areas of endemicity. Recently, at least three HEV strains have been isolated in the United States from three human patients with acute hepatitis E (5, 11, 21, 28). The sequences of the viruses recovered from these patients appear to diverge substantially from those of all other known human HEVs. Subsequently, additional HEVs have been identified from hepatitis E patients in countries where HEV is not endemic (Greece [22], Italy [37], Spain [19], Japan [25], Austria [32], and Argentina [23]), and accordingly, their partial sequences have become available. Limited available information suggests that these viruses form a genetically separate group distinct from the two known classical types of HEV.

In parallel with these studies, an HEV strain has been isolated from pigs in the United States (16). This swine form of HEV is divergent from the two classical types of human HEV, with sequence identities of only 74%, but it appears to be rather similar to newly isolated human HEV strains US-1 and US-2, with 91% sequence identities. The swine form of HEV has subsequently been identified in pigs in other countries such as Taiwan, Spain, Japan, and New Zealand (6, 7, 8, 18, 19). Accordingly, our initial interest was to examine Canadian swine herds for the prevalence of antibodies to HEV. Our study indicated that a high percentage (up to 80%) of pigs on Ontario farms became seropositive for antibodies to HEV by 6 months of age (36). Subsequent identification of the viral se-

* Corresponding author. Mailing address: Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Phone: (519) 824-4120, ext. 4729. Fax: (519) 767-0809. E-mail: dyoo@uoguelph.ca.

quence from feces and sequencing of a small region in open reading frame (ORF) ORF2, which represents nucleotide positions 5620 to 5880 with respect to the genome of the U.S. swine HEV isolate (strain US swine), showed that the sequence identity between the Canadian swine HEV strains isolated in the province of Saskatchewan and HEV strain US swine was only 85.8% (36). This suggested that HEV strains circulating in pigs in Canada might represent a genetically distinct branch of swine HEV. The limited available partial sequence information of these swine HEV strains isolated from pigs in different geographic regions suggests that the swine HEV strains may be genetically divergent from the prototype swine HEV strain isolated in the United States, strain US swine (36). In these regards, we examined an HEV isolate from pigs in Canada and determined its full-length genomic sequence. Our study reveals that the swine HEV strain isolated in Canada is significantly divergent from other known swine HEV strains, confirming that genetic variations exist among swine HEV strains.

MATERIALS AND METHODS

Fecal specimens and extraction of RNA. Fecal specimens were collected from 3- to 4-month-old pigs housed in the Arkell Research Station of the University of Guelph, Guelph, Ontario, Canada. The fecal materials were diluted to a 10% (wt/vol) suspension in phosphate-buffered saline, and the suspensions were centrifuged in a bench-top clinical centrifuge (model TJ-6; Beckman Instruments) at 3,000 rpm for 10 min at room temperature. The supernatant was collected and aliquoted for storage at −70°C until use. Viral RNA was extracted from a 100-μl aliquot of the 10% fecal suspension by using TriZol (Invitrogen Canada Inc., Mississauga, Ontario, Canada) according to the instructions of the manufacturer, and the entire amount of RNA was subjected to cDNA synthesis.

Reverse transcription reaction and PCR. The oligonucleotide primers used for synthesis of cDNA and PCR amplification were purchased from Sigma Genosys (Austin, Tex.). The first strand of cDNA was synthesized at 42°C for 2 h with 200 U of Superscript II RNase H-reverse transcriptase (Invitrogen) in the presence of 10% dimethyl sulfoxide (DMSO). A 30-μl PCR mixture with 3 μl of the synthesized cDNA and an external set of forward and reverse primers was prepared. The first-round PCR was carried out for 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 3 min. At the end of the first-round PCR, 3 μl was taken to prepare a 30-μl reaction mixture for the second-round PCR with an internal primer set. The second-round PCR was carried out for an additional 35 cycles of denaturation at 94°C for 45 s, annealing at 55°C for 45 s, and extension at 72°C for 2 min. The nucleotide sequences at the 5′ and 3′ termini of the genome were determined with the 5′ RACE and 3′ RACE kit (Invitrogen) according to the instructions of the manufacturer. To clone the hypervariable region in ORF1 and the region in ORF3 with high G+C contents, PCR was performed in the presence of 10% (vol/vol) DMSO and/or by using 0.5 M GC-Melt (Clontech Laboratories Inc., Palo Alto, Calif.). The PCR conditions for amplification of the different regions were slightly variable, depending on the success of the amplification and optimization of the conditions. For sequencing, the amplified product was purified by use of the GeneClean kit (Bio 101, San Diego, Calif.) and was sequenced directly or cloned into the pTrueBlue vector (Bio/Can Scientific, Mississauga, Ontario, Canada), followed by sequencing of the plasmid DNA. Nucleotide sequences were determined by PCR-based automatic dideoxy cycle sequencing (Perkin-Elmer, Norwalk, Conn.) at the Guelph Molecular Supercentre, University of Guelph.

Sequence analysis. Nucleotide sequences were assembled and analyzed with the OMIGA software program (version 2.0; Oxford Molecular Inc., Madison, Wis.) and VectorNTI software (Informax, Inc., Gaithersburg, Md.). Phylogenetic trees were constructed by maximum parsimony methods with the aid of the PAUP software package (version 3.1.1) by using the unrooted dendrogram option. Confidence for the grouping in the trees was assessed by the bootstrap method (100 replicates). The consensus trees were viewed by using the TREEVIEW program. The sequences of the following strains were retrieved from GenBank (the accession numbers are given in parentheses) and were used for comparative analysis of HEV genomic sequences: Argentina-1 (AF264009), Argentina-2 (AF264010), Austria (AF279122), Burma (M73218), Canada swine SK3 (AF347692), China-1 (L25547), China-2 (AF082093), China-15 (M94177), New Chinese (AJ272108), Greece-1 (AF110391), Greece-2 (AF110392), India-1 (X99441), India-2 (X98292), Italy (AF110387), Japan (AP003430), Japan swine (AB073912), Mexico (M74506), Netherlands swine 15 (AF336000 and AF332620), Netherlands swine-22 (AF336291 and AF336002), Netherlands swine-50 (AY032758 and AF336006), Netherlands swine-105 (AF336298 and AF336013), Netherlands swine-36 (AF336293 and AF336004), Netherlands swine-82 (AF336009), Netherlands swine-85 (AF336295 and AF336010), Netherlands swine-122 (AF336299 and AF336014), New Zealand swine (AF200704 and AF215661), Spain-1 (AF195061 and AF195064), Spain-2 (AF195062 and AF195065), Spain swine (AF195063), Pakistan (M80581), Taiwan-1 (AF117275), Taiwan-2 (AF117277), Taiwan-3 (AF117279), Taiwan swine-1 (AF117280), Taiwan swine-2 (AF117281), US-1 (AF060668), US-2 (AF060669), US swine (AF082843).

Nucleotide sequence accession number. The full-length genomic sequence of HEV strain Arkell reported in this study has been deposited in the GenBank database under accession number AY115488.

RESULTS

Full-length genomic sequence of swine HEV strain Arkell. Previous studies have suggested that a genetic heterogeneity might exist in swine HEV strains, as is observed in human HEV strains. To examine the possible genetic variations in swine HEV strains, we attempted to determine the full-length genomic sequence of HEV using a selected HEV isolate. We recovered a virus from the feces of pigs housed at the Arkell Research Station of the University of Guelph and named it strain Arkell. The pigs housed at the Arkell Research Station were considered specific pathogen free. Using a total of 30 sets of primers and two rounds of consecutive PCRs for each set of primers, we were able to obtain the full-length genomic sequence of the Arkell swine HEV strain. The full-length genomic sequence of strain Arkell was determined to be 7,242 nucleotides, excluding the polyadenylated tail at the 3′ terminus of the genome. The organization of the viral genome consisted of a 5′ untranslated region (UTR) at nucleotide positions 1 to 26, ORF1 at positions 27 to 5150 in the +3 frame, ORF3 at positions 5150 to 5515 in the +2 frame, ORF2 at positions 5188 to 7167 in the +1 frame, and a 3′ UTR at positions 7168 to 7242, followed by a poly(A) tail of at least 15 A residues.

(i) 5′ UTR. The 5′ UTR was relatively short, with only 26 nucleotides. The 5′-terminal sequence of strain Arkell was identical in length to that of HEV strain US swine, but it was 9 nucleotides shorter than that of human HEV strain US-2. Among all known HEV strains of human and swine origin, strain US-2 is the only HEV strain that contains the extra 9 nucleotides, and the significance of this extra sequence remains unknown. The HEV genome originating from the swine species, including the Arkell virus, started with a G residue, whereas most other HEV strains of human origin contained an extra A nucleotide in front of the G, and thus, A was the first nucleotide of the genome.

(ii) ORF1. ORF1 of the Arkell virus consisted of 5,124 nucleotides capable of coding for a protein of 1,708 amino acids. When the Arkell virus sequence was aligned to the sequences of other HEV isolates, the Arkell virus sequence was found to be identical in length to that of HEV strain US swine. During the course of cloning and sequencing, the hypervariable region of ORF1 and the overlapping region of ORFs 2 and 3 tended to generate inconsistent sequencing results. To eliminate the possibility of a cloning artifact, we

attempted various protocols, which included an increase in the reaction temperature to 50°C for reverse transcription, addition of DMSO, and the use of GC-Melt (Clontech Laboratories Inc.). GC-Melt has successfully been used for the cloning and sequencing of GC-rich regions in HEV (31). Heating of the mixture of the RNA template and primers to 95°C with or without 30% DMSO prior to the reverse transcription reaction did not improve the conditions that possibly resulted in a cloning artifact, but inclusion of 0.5 M GC-Melt resulted in a larger fragment, regardless of whether the disassociation temperature was 70 or 95°C or whether 30% DMSO was absent or present. Under our defined conditions, a fragment of identical size was consistently produced, and each fragment from individual reactions was sequenced and identical results were obtained. On the basis of these observations, we concluded that HEV strain Arkell contained no major deletion or addition in the hypervariable region in comparison with HEV isolate US swine. A triplet codon of GTT (glutamic acid) present in the HEV strain US swine sequence at nucleotide positions 2193 to 2195 was found to be deleted in the Arkell strain. Instead, another triplet codon of CCG (serine) was inserted at positions 2361 to 2363 in the Arkell virus, resulting in an ORF1 sequence of identical size between strain US swine and the Canadian swine HEV isolate (Fig. 1). The biological implication of the hypervariable region remains unknown.

Pairwise comparisons of the full-length genomic sequence of the Arkell virus and those of HEV strain US swine, HEV strain Japan swine, and the two U.S. human HEV isolates indicated that sequence identities fell within the range of 85 to 87% (Table 1). The identities of the Arkell virus sequence with the sequences of the genotype 1 (Asian type) and genotype 2 (Mexican type) human HEV isolates were much lower (70 to 72%) (Table 1). At the amino acid level, the similarities of the ORF1 amino acid sequence of the Arkell virus with those of genotypes 1 and 2 were 80 to 81%, and the amino acid sequence similarities with those of isolates US swine and Japan swine were 95 to 96%. The motifs predicted within the ORF1 protein were all conserved in the Arkell virus, which included the putative nucleoside triphosphate-binding domain (GVPGSGKS), the DEAP motif in the putative helicase region, and the GDD motif in the RNA-dependent RNA polymerase.

(iii) ORF2. ORF2 of the Arkell virus consisted of 1,980 nucleotides and was predicted to code for a protein of 660 amino acids. This was identical in size to those of the U.S. human HEV and Burmese human HEV isolates but was 1 amino acid larger than that of the Mexican strain (which is 659 amino acids). When the Arkell strain ORF2 amino acid sequence was compared to those of three newly isolated Chinese HEVs (7, 8, 31), the ORF2 sequence was 12 amino acids smaller and 8 amino acids larger, depending on the frames of the isolates (672 and 652 amino acids, respectively). Pairwise comparisons showed that the Canadian Arkell virus had only 88 to 89% and 78 to 80% nucleotide sequence identities with the U.S. human HEV isolates and other known classical types of HEVs, respectively (Table 2). The similarities appeared to be much higher at the amino acid sequence level (97 to 98% and 91%, respectively). For ORF2, changes occurred mostly in the third positions of codons, resulting in 97% amino acid sequence identity. In comparison with HEV strain US swine,

```
Can-swine  (686) HIWESANPFCGEGTLYTRTWSTSGFSSDFSPP-EAAALPAAAPGPLPPTPVSDIWVPSLSGESQVDAGLVPPAPEPARLPSPTVSTLPSPPPPPVRKPPAPPLPRTR
US-swine   (686) ....P.....................EA..PVL.....LPH.....L.PP.K.....AS.....G...SI.L...-..L.....T..PS..
US1        (?)   ....P.....................EA..PAM..T..LPHS.....L.PP.E.F....AP.....D..G..G.V.L.P.-.....-.H..SI..PS.N.
US2        (686) .....P.....................EA..PAS.....LPY.....L.PP.E..H...AS..SV....G.T..I.L.P.-.....ATS..P....
Jap-swine  (686) .....P.....................EA.DHA..ITS.LPHS...A....AL.PP.E.....MAP.....K.VG...SNEPIT.-----.....AS...
Japan      (686) .....P.....................EA..PA.....SSS.....L.P..E.P..G.PPA..T.....C..KPNT.-----.....T...PS...
Mexico     (686) .E.R.....S.....ITDTPLTVG-.LISGHLD...HSGP.PATVTGPA.GS.D.PDPDPLPDVTDGSRPSGARPAG-----------N.NGVPQ.
NewChi     (686) .S.....S.....V.....C...LEPCVPSMPPPAEVN--..V.L.ALPSEIMEPAQPPASEPAA.PSDSVDNSFSPT.SGAPIA..APAL.VTHLSGPR.
Burma      (686) .V.....S.....EVDAV.SPAR.DLGFMSEPSI.SRAATPTLAAPLPPPAPDDPSPPPSAPA.AE..SGATAGAPA----------ITHQTA.H.
China1     (686) .V.....S.....EVDAV.SPAQ.DLGFTSEPSI.SRAATLT.AAPLPPPAPDDPSPTPSAPARGE...GATAGAPA----------ITHQTA.H.
India1     (686) YV.....S.....S....S..EVDAV.SPAR.DLGFVSEPSI.SRAATPTLAAPLPPPAPDDPSPPPSAPAPDE..SGTTAGAPA----------ITHQMA.H.
```

FIG. 1. Amino acid sequence alignments of the hypervariable region in ORF1 of various human and swine HEV strains. Numbers in parentheses indicate the amino acid positions of the ORF1 protein. Only different amino acid residues are presented. Identical amino acids are indicated by dots, and deletions are indicated by dashes.

TABLE 1. Pairwise comparisons of the full-length nucleotide sequence and the nucleotide and deduced amino acid sequences of ORF1

Strain	% Identity[a]													
	Canadian swine	US swine	US-1	US-2	Japan swine	Japan	Mexico	New Chinese	Burma	China-1	China-2	India-1	India-2	Pakistan
Canadian swine		87	87	87	86	86	73	74	74	73	73	73	74	73
US swine	87 (96)		91	91	87	87	74	74	74	74	74	73	74	74
US-1	87 (95)	91 (97)		91	87	87	74	75	74	74	74	73	74	74
US-2	86 (95)	91 (97)	92 (97)		87	86	73	74	73	73	73	73	73	73
Japan swine	85 (95)	86 (97)	86 (96)	86 (96)		89	73	75	74	74	74	74	74	74
Japan	85 (95)	86 (96)	86 (96)	86 (96)	88 (96)		73	75	74	74	74	73	74	74
Mexico	72 (81)	73 (81)	71 (81)	72 (81)	72 (82)	72 (82)		73	75	75	75	74	75	74
New Chinese	71 (84)	73 (84)	72 (84)	73 (84)	73 (85)	73 (84)	72 (81)		74	74	74	74	74	74
Burma	71 (81)	72 (82)	71 (81)	71 (81)	71 (82)	72 (82)	73 (83)	73 (82)		93	93	96	93	93
China-1	71 (81)	72 (81)	71 (80)	72 (80)	72 (81)	72 (81)	73 (82)	72 (82)	93 (98)		98	92	94	98
China-2	70 (82)	72 (82)	71 (81)	72 (81)	71 (83)	72 (82)	73 (83)	72 (80)	93 (96)	98 (98)		91	93	98
India-1	71 (80)	71 (80)	71 (79)	71 (79)	71 (80)	71 (81)	72 (81)	72 (80)	95 (96)	91 (95)	91 (94)		92	91
India-2	71 (82)	73 (82)	72 (81)	72 (81)	71 (83)	72 (82)	73 (83)	72 (82)	93 (98)	93 (98)	93 (96)	91 (95)		93
Pakistan	71 (82)	72 (82)	72 (81)	72 (81)	72 (82)	72 (82)	73 (83)	72 (82)	93 (98)	99 (99)	98 (97)	91 (95)	93 (98)	

[a] The values represent the percent identities of the full-length nucleotide sequence (above the diagonal) or the ORF1 nucleotide and amino acid (in parentheses) sequences (below the diagonal).

most changes (11 of a total of 16 changes; 68%) appeared to occur in the overlapping region of ORFs 2 and 3. Although potentially hypervariable regions of 15 and 17 amino acids were identified when HEV strain US swine was compared to the prototype human HEV Burmese and Mexican isolates, respectively (9, 26), the Arkell virus sequence was different from the HEV strain US swine sequence by a single amino acid within this region.

(iv) ORF3. ORF3 of the Canadian Arkell virus was 366 nucleotides in length and was predicted to code for a protein of 122 amino acids. Compared with the ORF3 sequences of Asian and Mexican strains of HEV, both of which encoded a protein of 123 amino acids, isolates Arkell and US swine appeared to have a deletion of 3 nucleotides immediately following the AUG start codon. This caused one amino acid deletion in the ORF3 protein, resulting in a protein coding for 122 amino acids. The pairwise comparisons of the nucleotide and deduced amino acid sequences indicated that the ORF3 gene was more conserved than other genes at the nucleotide level but that it was more variable at the amino acid level. The Arkell virus shared 93 to 95% and 83 to 85% identities at the nucleotide level with the U.S. HEV strains (both human and swine strains) and all other known classical types of HEV isolates, respectively (Table 2). At the amino acid level, the identities were 94 to 96% and 77 to 81%, respectively (Table 2).

(v) 3′ UTR. The Arkell virus contained a 3′ UTR identical in length (75 nucleotides) to those of the U.S. (both human and swine) HEV isolates. This region was, however, variable in both length and sequence compared to those of other classical HEV isolates. There was an insertion of 7 nucleotides when the sequence was aligned to that of the Burmese isolate and a deletion of 2 nucleotides when the sequence was aligned to that of the Mexican isolate (Fig. 2). Comparisons of the 3′ UTR sequences indicated that the Arkell virus sequence shared 83 to 84% and 73% identities with those of the U.S. isolates (both human and swine) and the two classical types of HEV, respectively.

Phylogenetic analysis. The full-length genomic sequence of Canadian HEV strain Arkell was compared to those of other known HEV isolates including genotype 3 viruses. As of this

TABLE 2. Pairwise comparisons of the nucleotide and deduced amino acid sequences of ORFs 2 and 3

Strain	% Identity[a]													
	Canadian swine	US swine	US-1	US-2	Japan swine	Japan	Mexico	New Chinese	Burma	China-1	China-2	India-1	India-2	Pakistan
Canadian swine		88 (97)	89 (97)	89 (98)	89 (97)	89 (98)	78 (91)	79 (91)	80 (91)	79 (91)	79 (91)	79 (91)	79 (91)	79 (91)
US swine	94 (94)		92 (97)	92 (98)	88 (98)	88 (98)	78 (89)	80 (91)	79 (91)	79 (90)	80 (91)	79 (91)	79 (91)	80 (91)
US-1	93 (94)	94 (93)		91 (98)	88 (97)	87 (97)	78 (89)	79 (90)	79 (91)	78 (90)	79 (91)	79 (91)	78 (91)	79 (91)
US-2	95 (95)	97 (96)	95 (95)		89 (98)	88 (98)	79 (90)	80 (91)	79 (91)	79 (91)	79 (91)	79 (91)	79 (91)	79 (91)
Japan swine	94 (95)	95 (94)	93 (92)	96 (95)		91 (98)	77 (90)	80 (91)	80 (91)	80 (91)	80 (91)	80 (91)	80 (91)	80 (91)
Japan	96 (96)	97 (95)	94 (94)	97 (97)	97 (98)		78 (90)	80 (92)	80 (91)	80 (91)	80 (91)	79 (91)	80 (91)	79 (91)
Mexico	85 (77)	84 (81)	83 (83)	85 (83)	83 (78)	85 (79)		78 (88)	81 (93)	80 (92)	80 (92)	80 (92)	80 (92)	81 (93)
New Chinese	83 (77)	83 (75)	82 (79)	83 (75)	83 (75)	84 (75)	82 (73)		79 (89)	78 (89)	78 (89)	78 (89)	78 (89)	79 (89)
Burma	85 (80)	84 (81)	86 (83)	86 (83)	85 (81)	85 (82)	89 (100)	82 (73)		93 (98)	94 (99)	94 (98)	96 (99)	93 (99)
China-1	85 (80)	84 (80)	85 (83)	85 (81)	84 (80)	85 (81)	90 (98)	82 (72)	98 (98)		98 (96)	93 (98)	93 (98)	98 (99)
China-2	86 (81)	84 (80)	85 (81)	85 (81)	84 (81)	85 (81)	90 (98)	82 (74)	98 (98)	98 (96)		93 (98)	93 (98)	98 (99)
India-1	85 (81)	84 (81)	85 (83)	86 (83)	85 (81)	86 (82)	90 (98)	82 (73)	97 (98)	97 (96)	97 (96)		93 (98)	93 (98)
India-2	85 (81)	85 (81)	85 (83)	86 (83)	85 (81)	85 (82)	89 (100)	82 (73)	99 (100)	97 (98)	97 (98)	97 (98)		92 (99)
Pakistan	86 (81)	85 (81)	86 (83)	86 (83)	85 (81)	86 (82)	90 (100)	82 (73)	98 (100)	99 (98)	99 (98)	97 (98)	98 (100)	

[a] The values represent the percent identities of the nucleotide and amino acid (in parentheses) sequences for ORF2 (above the diagonal) and ORF3 (below the diagonal).

```
            Stop ** ***       ** *** **                          ******   *****      ***********   PolyA
Can-swine   TAA  TTAATTCTTTCT-GTGCCC-CCTTCGTAGATCTCTT---CTGCTATATTTCTTTTTTCTGCCTTTCGCGCTCCCTGG  AAAAA
Jap-swine   TAA  ......T.CCTB-......-.......G.T.....---TC.A.T.......C........TC.C.............  AAAAA
US-swine    TAA  .......C.TTTT.....-.....A...C.TC...---TG.T.T........A.......T..C.............  AAAAA
US1         TAA  .......C.TTT-......-......C..T.....---TG...T.......CA.......T..C.............  AAAAA
US2         TAA  .......C.T.T-......-.........T.TCT..---.C.T.T........A.......T..C.............  AAAAA
Japan       TAA  .......G.C.T-..A...-.........T.T....---T....T.......C........T..C.............  AAAAA
Burma       TAG  ..T...TGCTT--......-..C.T----C.T...G---T....-.......CA.......G..C...........-  AAAAA
Mexico      TAG  ..T...TGGCT--......A...A.T..T....GC.GATT.C..T......C.......CGG.CC...........-  AAAAA
```

FIG. 2. Nucleotide sequence alignments of the 3′ UTR of different 3 HEV strains. Stop represents the translation termination codon of ORF2. Asterisks indicate conserved nucleotides. Identical nucleotides are indicated by dots, and deletions are indicated by dashes.

writing, only two isolates of the swine form of HEV (one Japanese isolate and one U.S. isolate) have been fully sequenced (15, 16, 18); in addition, three human genotype 3 viruses (two U.S. isolates and one Japanese isolate) have been fully sequenced (5, 21, 25). However, small fragments representing different portions of the genomes of several additional isolates of genotype 3 viruses have been sequenced. We first compared the full-length genomic sequence of Canadian HEV strain Arkell to other full-length sequences of genotype 3 viruses (Fig. 3A). The Arkell HEV sequence was divergent from the two classical-type HEV sequences and, along with the U.S. and Japanese HEV isolates, formed a distinct branch among the HEV isolates. Strain Arkell differed from the other HEV isolates by at least 13 to 15%, suggesting the presence of genetic variations among genotype 3 viruses (Table 1). Having considered the geographic regions where the viruses were isolated, it was noteworthy that both the US-1 and the US-2 human HEV isolates were closer to HEV isolate US swine (5, 17, 21), and similarly, the Japanese human HEV isolate was closer to the Japanese swine HEV strain Japan swine than to other HEV isolates from distant geographic regions (Fig. 3A). The Canadian Arkell virus appeared to be distant from HEV strains isolated from the United States or Japan. Although several HEV sequences from pigs have been available in some countries to date, only partial sequences have been determined. Furthermore, these partial sequences represent different regions of the genome; therefore, it is not possible to compare their sequences directly to each other. Despite these limitations, however, we constructed two additional phylogenetic trees for genotype 3 HEV isolates by using the 242-nucleotide region in ORF1 that encodes the methyltransferase gene and the 304-nucleotide region in ORF2 that encodes the ORF2 capsid protein gene (Fig. 3B and C). These trees also suggest a possible geographic relatedness between human and swine HEV isolates belonging to genotype 3. While further information is required to make the observation valid, it is tempting to speculate that the HEV strains identified in humans in certain areas may have been derived from HEV strains circulating in pigs in the same geographic area, as it was observed from the phylogenetic tree that the U.S. human HEV isolates were closer to the U.S. swine HEV isolate, the Taiwan human HEV isolates were closer to the Taiwan swine HEV isolates, and the Japan human HEV isolates were closer to the Japanese swine HEV isolate. Similarly, the European swine HEV isolates were closer to the European human HEV isolates, with some exceptions. This observation supports the hypothesis that swine HEV isolates may be zoonotic agents from pigs infecting humans.

DISCUSSION

Clinical cases of hepatitis E have been rare in Western Europe, Oceania, and North American countries; and the cases in these countries are often associated with travel to regions where hepatitis E is endemic. Recently, HEV has been identified in human hepatitis E patients in the United States, where HEV is considered nonendemic, with no association with international travel (13, 21). Subsequently, the HEV sequence has been identified in humans in Austria, Italy, Spain, Greece, and Japan (19, 25, 32, 37). When the sequences of these isolates are compared to those of two classical types of HEV (designated genotype 1 and genotype 2 for the Asian and Mexican genotypes of HEV isolates, respectively) (31), the sequence identities are strikingly low (only 70 to 72%). In parallel, an HEV strain has been isolated from pigs in the United States (22), and it has been found that the nucleotide sequence of this swine form of HEV is similar to the sequences of the human HEV strains isolated from countries where HEV is believed to be nonendemic. The U.S. human HEV strains and the HEV strain US swine, all of which were isolated in the United States, where HEV is considered nonendemic, show an overall sequence identity of 91%, whereas the identities of the nucleotide sequences of these strains compared to those of the genotype 1 and 2 strains are 73 to 74% (Table 1).

No known clinical case of HEV in humans has been reported in Canada, and accordingly, Canada has been considered free of HEV, at least in humans. Recent studies, however, have revealed that a large proportion (38 to 88%, depending on the areas tested) of commercial pigs become seropositive for HEV by 6 months of age (36), suggesting active circulation of HEV in Canadian swine herds. Indeed, a portion of the HEV genomic sequence has been identified from the feces of pigs, and the virus has been recovered from these pigs. This piece of evidence demonstrates that HEV is truly circulating in pigs and the virus is being shed in the feces of infected animals. The HEV-infected pigs are clinically normal and do not exhibit any abnormal symptoms as a result of the infection. Similar serological observations have been made for pigs in the United States, New Zealand, and Australia (3, 6, 22). All of these countries are considered free of human HEV, but swine HEV appears to be prevalent. A recent study with a relatively large number of subjects (389 individuals in the United States who work with swine) indicates that swine veterinarians are at higher risk of HEV infection than healthy blood donors (17). Independently from this study, swine farmers have also been shown to have significantly higher levels of antibodies to HEV (14). Thus, it seems that HEV in pigs may cross species bar-

riers and be able to infect humans. HEV appears to be ubiquitous in pig populations worldwide, as it has been demonstrated in the United States, Canada, New Zealand, Australia, Taiwan, Korea, China, Japan, Spain, India, and The Netherlands (1, 3, 6, 10, 14, 18, 19, 30, 36). Taking these facts together, HEV is likely a zoonotic agent that is transmitted from pigs to humans.

Except for the genomes of the prototype swine HEV strain isolated in the United States and the swine HEV strain recently isolated in Japan, whose full-length genomic sequences have been determined, only small portions of the genomes from other swine HEV isolates have been sequenced to date; and the remainder of the genome is largely undetermined. Accordingly, little is known about the genomic nature of the swine form of HEV, while many strains of human HEV have been completely sequenced. Sequencing of the HEV genome is cumbersome. Difficulties include a lack of a virus cultivation system, a relatively short period of a viremic state in vivo upon infection, a low level of virus shedding in the feces, low degrees of sequence homology among HEV isolates, and high G+C contents in some parts of the genome. Naturally infected animals are healthy and have subclinical infections; therefore, it is difficult to identify HEV-positive animals in order to collect appropriate samples for evaluation. Despite these obstacles, we have determined the full-length genomic sequence of a swine HEV isolate directly from fecal specimens. The Arkell strain of HEV isolated from pigs in Ontario, Canada, appears to be genetically distant from the two classical types of HEV, with sequence identities to the genotype 1 and 2 sequences of only 73 and 74%, respectively. Instead, the Canadian Arkell virus sequence is rather closely related to that of the prototype swine HEV strain isolated in the United States and the swine HEV strain recently isolated in Japan, and the sequences of these isolates are some distance from those of both classical types (Fig. 3A). When the full-length Arkell virus genomic sequence is compared with those of genotype 3 HEV isolates, the sequence identities appear to be much lower than expected (only 87%). The phylogenetic studies indicate that the Canadian Arkell virus represents a distinct variant among the genotype 3 isolates (Fig. 3A). We have compared the sequences of the individual ORFs of the Arkell virus with the respective sequences of the ORFs of other swine HEVs, and very similar phylogenetic observations have been obtained (Fig. 3B and C). Partial sequences of swine HEV isolates have additionally been identified in New Zealand and Spain (6, 19). Phylogenetic studies with the available partial sequence of the ORF1 region indicate that the New Zealand swine HEV isolate is closely related to the European (Italy) human HEV isolate (88% identity) (Fig. 3B) (32, 37). When the New Zealand swine HEV isolate is compared to the Canadian strain, it is found that they share 85% identity. The Spanish swine HEV strain has 83.9% similarity to the New Zealand swine HEV in a region of 304 nucleotides in ORF2. The Spanish human HEV isolate and two Spanish swine HEV isolates share similarities of 92.1 and 94.0%, suggesting a possible relatedness of the swine and human HEV isolates in the same geographic region. While further information is required, it seems that genetic variations exist within genotype 3 HEV strains and that swine HEV isolates may have infected humans in the same geographic region. A recent study with a collection of swine HEV isolates recovered from six states in the United States also demonstrates the sequence heterogeneity of swine HEV isolates (10, 24).

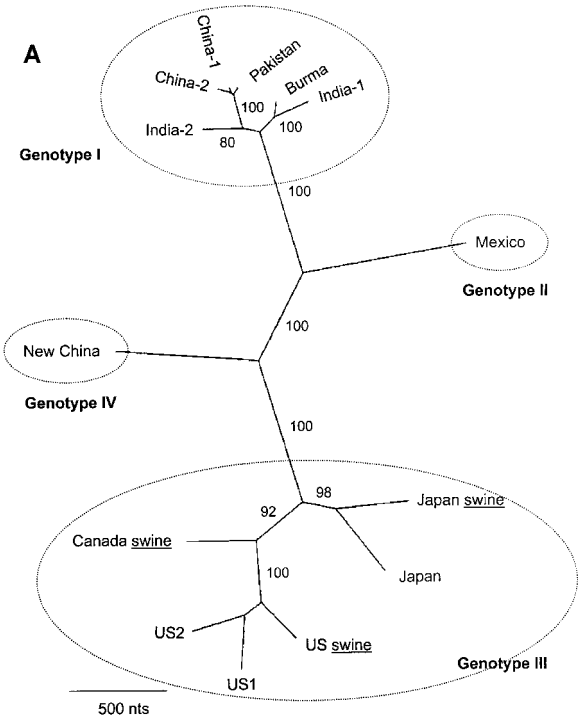

FIG. 3. Unrooted phylogenetic trees depicting the relationships of the genomic sequences of swine and human HEV strains. (A) Full-length genomic sequence; (B) 242 nucleotides encoding the nonstructural methyltransferase gene in ORF1; (C) 304 nucleotides encoding the ORF2 structural gene. HEV strains isolated from pigs are underlined. The scales represent the number of nucleotide changes (nts). The internal node numbers represent the bootstrap values as a percentage of trees obtained from 100 replicates. The GenBank accession numbers for each sequence are presented in Materials and Methods.

Canadian swine HEV strain Arkell is distantly related from HEV strain US swine, while HEV strain US swine is more closely related to the U.S. human HEV isolates. In Canada, approximately 2% of the general population and 2.4% of Canadian external affairs employees have been found to be seropositive for HEV (4, 35), and a higher rate (7.4%) of positivity for antibodies to HEV has been shown among Indochinese immigrants. These cases are likely associated with travel to and from areas where HEV is endemic. Despite the absence of clear evidence of the active circulation of a human form of HEV in Canada, it is tempting to speculate that if HEV strains were isolated from humans residing in Ontario, Canada, they would be more closely related to the Arkell HEV strain than to the U.S. HEV strains. Recently, several new swine and human HEV sequences have been identified in China, where HEV is endemic, and in Taiwan, where HEV is not endemic. Sequence comparisons indicate that these HEV isolates are distant from the classical type of HEV but share rather high degrees of similarity to each other. The newly isolated Chinese HEV strain appears to be distinct from all three known genotypes and has been suggested to be the fourth genotype (31). Accordingly, swine HEV strains isolated

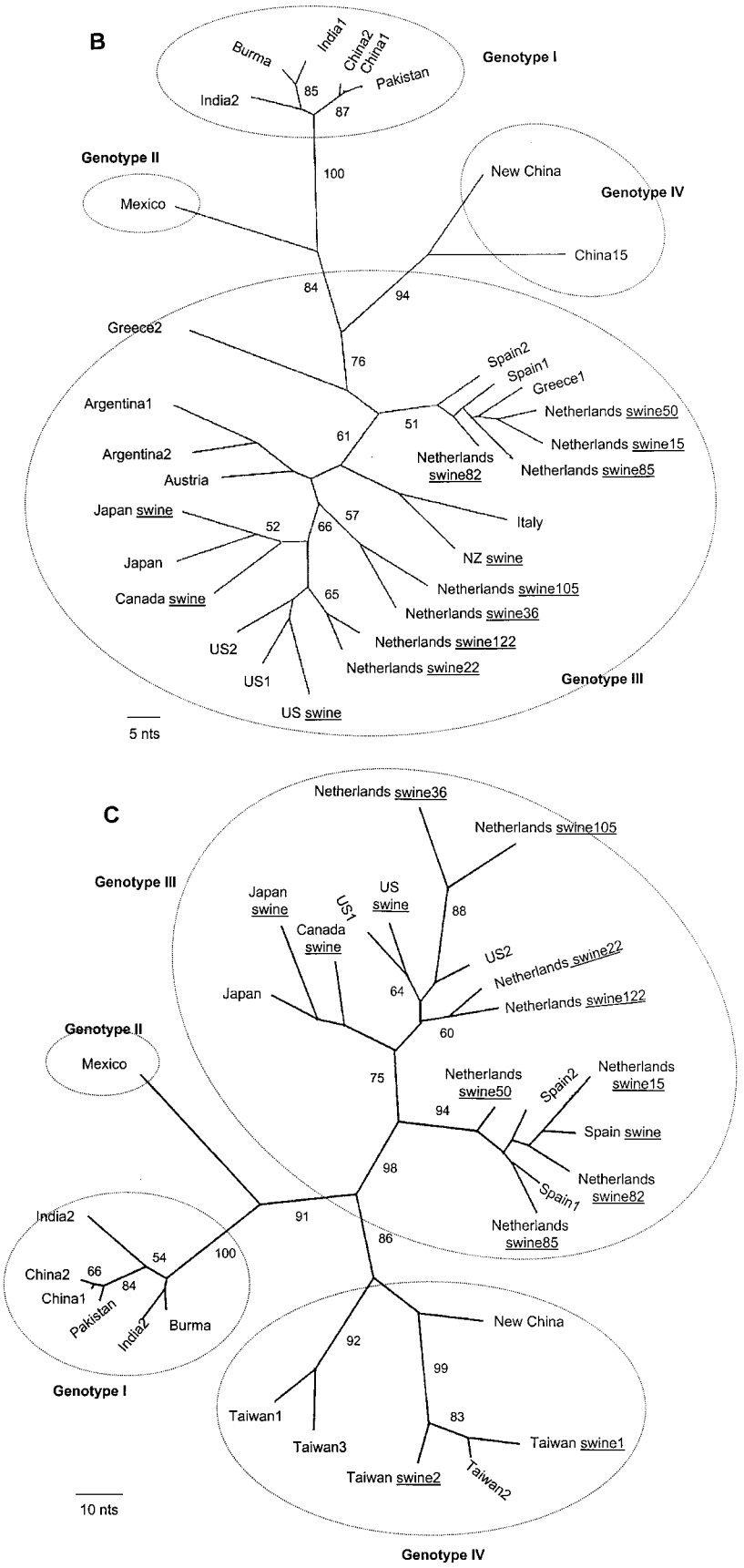

FIG. 3—*Continued.*

in Taiwan are clustered within genotype 4, while swine HEV strains isolated from imported U.S. pigs remain genotype 3 (33, 34). In contrast, swine HEV strains in India have been identified to be genotype 4, whereas human HEV strains in the same region remain genotype 1 (1). Further studies are required to verify the geographic relationship between human and swine HEV strains.

Pigs naturally infected with swine HEV appear to be asymptomatic and healthy. Since a large proportion of the human population is reported to be seropositive for HEV and most of these individuals are asymptomatic, swine HEV infection seems mild. Swine is considered an attractive donor animal species for xenotransplantation, and in this regard, swine HEV is of concern as a potential xenogeneic agent that can be transmitted from pigs to humans via organ transplantation, especially under immunosuppressive conditions (35). All other known swine HEV isolates reported to date have been identified in commercial swine herds, and our study is the first report on the isolation of swine HEV from specific-pathogen-free pigs. Our study indicates that swine HEV may be more widely distributed in pigs than was previously thought and that it is distributed not only in pig production units but also in animals raised in a closed environment for specific purposes. It raises a concern that potential donor pigs for organ transplantation should be closely monitored for the presence of swine HEV, even if they are housed under isolated conditions. The sensitivities and the specificities of the enzyme-linked immunosorbent assay and reverse transcription-PCR used for detection of HEV may need to be further elaborated, and the present study may be helpful in improving such techniques.

ACKNOWLEDGMENTS

This study was supported by funding from Imutran Novartis (now Immerge Biotherapeutics) and Health Canada (Bureau of Infectious Diseases Bloodborne Pathogens Division).

We are grateful to Jinzhong Fu for help with phylogenetic tree constructions and sequence analyses and to Grace Yoo for editorial assistance.

REFERENCES

1. **Arankall, V. A., L. P. Chobe, M. V. Joshi, M. S. Chadha, B. Kundu, and A. M. Walimbe.** 2002. Human and swine hepatitis E viruses from western India belong to different genotypes. J. Hepatol. **36:**417–425.
2. **Berke, T., and D. O. Mastson.** 2000. Reclassification of the Caliciviridae into distinct genera and exclusion of hepatitis A virus from the family on the basis of comparative phylogenetic analysis. Arch. Virol. **145:**1421–1436.
3. **Chandler, J. D., M. A. Riddell, F. Li, R. J. Love, and D. A. Anderson.** 1999. Serological evidence for swine hepatitis E virus infection in Australian pig herds. Vet. Microbiol. **68:**95–105.
4. **Chaudhary, R. K., M. Morbey, and P. O. Yarbough.** 1995. Detection of hepatitis E virus antibody in Canada. Can. J. Infect. Dis. **6:**264–265.
5. **Erker, J. C., S. M. Desai, G. G. Schlauder, G. J. Dawson, and I. K. Mushahwar.** 1999. A hepatitis E virus variant from the United States: molecular characterization and transmission in *Cynomolgus* macaques. J. Gen. Virol. **80:**681–690.
6. **Garkavendo, O., A. Obfiadina, X. J. Meng, D. A. Anderson, H. J. Bemard, B. A. Schroeder, Y. E. Khudyakov, and M. C. Croxson.** 2001. Detection and characterization of swine hepatitis E virus in New Zealand. J. Med. Virol. **65:**525–529.
7. **Hsieh, S. Y., P. Y. Yang, Y. P. Ho, C. M. Chu, and Y. F. Liaw.** 1998. Identification of a novel strain of hepatitis E virus responsible for sporadic acute hepatitis in Taiwan. J. Med. Virol. **55:**300–303.
8. **Hsieh, S. Y., X. J. Meng, Y. H. Wu, S. T. Liu, A. W. Tam, D. Y. Lin, and Y. F. Liaw.** 1999. Identity of a novel swine hepatitis E virus in Taiwan forming a monophyletic group with Taiwan isolates of human hepatitis E virus. J. Clin. Microbiol. **37:**3828–3834.
9. **Huang, C. C., D. Nguyen, J. Fernandez, K. Y. Yun, K. E. Fry, D. W. Bradley, A. W. Tam, and G. R. Reyes.** 1992. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). Virology **191:**550–558.
10. **Huang, F. F., G. Haqshenas, D. K. Guenette, P. G. Halbur, S. K. Schommer, F. W. Pierson, T. E. Toth, and X. J. Meng.** 2002. Detection by reverse transcription-PCR and genetic characterization of field isolates of swine hepatitis E virus from pigs in different geographic regions of the United States. J. Clin. Microbiol. **40:**1326–1332.
11. **Kabrane-Lazizi, Y., M. Zhang, R. H. Purcell, K. D. Miller, R. T. Davey, and S. U. Emerson.** 2001. Acute hepatitis caused by a novel strain of hepatitis E virus most closely related to United States strains. J. Gen. Virol. **82:**1687–1693.
12. **Kumar, R. M., S. Uduman, S. Rana, J. K. Kochiyil, A. Usmani, and L. Thomas.** 2001. Sero-prevalence and mother-to-infant transmission of hepatitis E virus among pregnant women in the United Arab Emirates. Eur. J. Obstet. Gynecol. Reprod. Biol. **100:**9–15.
13. **Kwo, P. Y., G. G. Schlauder, H. A. Carpenter, P. J. Murphy, J. E. Rosenblatt, G. J. Dawson, E. E. Mast, K. Krawczynski, and V. Balan.** 1997. Acute hepatitis by a new isolate acquired in the United States. Mayo Clin. Proc. **72:**1133–1136.
14. **Meng, X. J., S. Dea, R. E. Engle, R. M. Friendship, Y. S. Lyoo, T. Sirinarumitr, K. Urairong, D. Wang, D. Yoo, Y. Zhang, R. H. Purcell, and S. U. Emerson.** 1999. Prevalence of antibodies to the hepatitis E virus in pigs from countries where hepatitis E is common or is rare in the human population. J. Med. Virol. **59:**297–302.
15. **Meng, X. J., P. G. Halbur, M. S. Shapiro, S. Govindarajan, J. D. Bruna, I. K. Mushahwar, R. H. Purcell, and S. U. Emerson.** 1998. Genetic and experimental evidence for cross-species infection by swine hepatitis E virus. J. Virol. **72:**9714–9721.
16. **Meng, X. J., R. H. Purcell, P. G. Halbur, J. R. Lehman, D. M. Webb, T. S. Tsareva, J. S. Haynes, B. J. Thacker, and S. U. Emerson.** 1997. A novel virus in swine is closely related to the human hepatitis E virus. Proc. Natl. Acad. Sci. USA **94:**9860–9865.
17. **Meng, X. J., B. Wiseman, F. Elvinger, D. K. Guenette, T. E. Toth, R. E. Engle, S. U. Emerson, and R. H. Purcell.** 2002. Prevalence of antibodies to hepatitis E virus in veterinarians working with swine and in normal blood donors in the United States and other countries. J. Clin. Microbiol. **40:**117–122.
18. **Okamoto, H., M. Takahashi, T. Nishizawa, K. Fukai, U. Muramatsu, and A. Yoshikawa.** 2001. Analysis of the complete genome of indigenous swine hepatitis E virus isolated in Japan. Biochem. Biophys. Res. Commun. **289:**926–936.
19. **Pina, S., M. Buti, M. Cotrina, J. Piella, and R. Girones.** 2000. HEV identified in serum from humans with acute hepatitis and in sewage of animal origin in Spain. J. Hepatol. **33:**826–833.
20. **Purcell, R. H.** 1996. Hepatitis E virus, p. 2831–2832. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields virology, 3rd ed. Lippincott-Raven Publishers, Philadelphia, Pa.
21. **Schlauder, G. G., G. J. Dawson, J. C. Erker, P. Y. Kwo, M. F. Knigge, D. L. Smalley, J. E. Rosenblatt, S. M. Desai, and I. K. Mushahwar.** 1998. The sequence and phylogenetic analysis of novel hepatitis E virus isolated from a patient with acute hepatitis reported in the United States. J. Gen. Virol. **79:**447–456.
22. **Schlauder, G. G., S. M. Desai, A. R. Zanetti, N. C. Tassopoulos, and I. K. Mushahwar.** 1999. Novel hepatitis E virus (HEV) isolates from Europe: evidence for additional genotypes of HEV. J. Med. Virol. **57:**243–251.
23. **Schlauder, G. G., S. S. Frider, G. C. Castano, and I. K. Mushahwar.** 2000. Identification of 2 novel isolates of hepatitis E virus in Argentina. J. Infect. Dis. **182:**294–297.
24. **Schlauder, G. G., and I. K. Mushahwar.** 2001. Genetic heterogeniety of hepatitis E virus. J. Med. Virol. **65:**282–292.
25. **Takahashi, K., K. Iwate, N. Watanabe, T. Hatahara, Y. Ohta, K. Baba, and S. Mishiro.** 2001. Full-genome nucleotide sequence of a hepatitis E virus strain that may be indigenous to Japan. Virology **287:**9–12.
26. **Tam, A. W., M. M. Smith, M. E. Guerra, C. C. Huang, D. W. Bradley, K. E. Fry, and G. R. Reyes.** 1991. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. Virology **185:**120–131.
27. **Ticehurst, J. R.** 1999. Hepatitis E virus, p. 1053–1069. *In* P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Yolken (ed.), Manual of clinical microbiology, 7th ed. American Society for Microbiology, Washington, D.C.
28. **Tsang, T. H., E. K. Denison, H. V. William, L. V. Venczel, M. M. Ginsberg, and D. J. Vugia.** 2000. Acute hepatitis E infection acquired in California. Clin. Infect. Dis. **30:**618–619.
29. **Tsarev, S. A., S. U. Emerson, G. R. Reyes, T. S. Tsareva, L. J. Legters, I. A. Malik, and R. H. Purcell.** 1992. Characterization of a prototype strain of hepatitis E virus. Proc. Natl. Acad. Sci. USA **89:**559–563.
30. **Van der Poel, W. H. M., F. Verschoor, R. van der Heidi, M. I. Herrera, A. Vivo, M. Kooreman, and A. M. de Roda Husman.** 2001. Hepatitis E virus sequences in swine related to sequenced in humans, The Netherlands. Emerg. Infect. Dis. **7:**970–976.
31. **Wang, Y., H. Zhang, R. Ling, H. Li, and T. J. Harrison.** 2000. The complete sequence of hepatitis E virus genotype 4 reveals an alternative strategy for

translation of open reading frame 2 and 3. J. Gen. Virol. **81:**1675–1686.
32. **Worm, H. C., G. G. Schlauder, H. Wurzer, and I. K. Mushahwar.** 2000. Identification of a novel variant of hepatitis E virus in Austria: sequence, phylogenetic and serological analysis. J. Gen. Virol. **81:**2885–2890.
33. **Wu, J. C., C. M. Chen, T. Y. Chiang, I. J. Sheen, J. Y. Chen, W. H. Tsai, Y. H. Hwang, and S. D. Lee.** 2000. Clinical and epidemiological implications of swine hepatitis E virus infection. J. Med. Virol. **60:**166–171.
34. **Wu, J. C., C. M. Chen, T. Y. Chiang, W. H. Tsai, W. J. Jeng, I. J. Sheen, C. C. Lin, and X. J. Meng.** 2002. Spread of hepatitis E virus among different-age pigs: two-year survey in Taiwan. J. Med. Virol. **66:**488–492.
35. **Yoo, D., and A. Giulivi.** 2000. Xenotransplantation and the potential risk of xenogeneic transmission of porcine viruses. Can. J. Vet. Res. **64:**193–203.
36. **Yoo, D., P. Willson, Y. Pei, M. A. Hayes, A. Deckert, C. E. Dewey, R. M. Friendship, Y. Yoon, M. Gottschalk, C. Yason, and A. Giulivi.** 2001. Prevalence of hepatitis E virus antibodies in Canadian swine herds and identification of a novel variant of swine hepatitis E virus. Clin. Diagn. Lab. Immunol. **8:**1213–1219.
37. **Zanetti, A. R., G. G. Schlauder, L. Romano, E. Tanzi, P. Fabris, G. J. Dawson, and I. K. Mushahwar.** 1999. Identification of a novel variant of hepatitis E virus in Italy. J. Med. Virol. **57:**356–360.